Neuronal cell-type classification: challenges, opportunities and the path forward

Hongkui Zeng[1] and Joshua R. Sanes[2]

Abstract| Neurons have diverse molecular, morphological, connectional and functional properties. We believe that the only realistic way to manage this complexity — and thereby pave the way for understanding the structure, function and development of brain circuits — is to group neurons into types, which can then be analysed systematically and reproducibly. However, neuronal classification has been challenging both technically and conceptually. New high-throughput methods have created opportunities to address the technical challenges associated with neuronal classification by collecting comprehensive information about individual cells. Nonetheless, conceptual difficulties persist. Borrowing from the field of species taxonomy, we propose principles to be followed in the cell-type classification effort, including the incorporation of multiple, quantitative features as criteria, the use of discontinuous variation to define types and the creation of a hierarchical system to represent relationships between cells. We review the progress of classifying cell types in the retina and cerebral cortex and propose a staged approach for moving forward with a systematic cell-type classification in the nervous system.

The founding document of neurobiology is Ramón y Cajal's 'Histology of the Nervous System of Man and Vertebrates' (REF. 1), which is perhaps the only 100-year-old manuscript that is still routinely consulted by neuroscientists. The work was, to a large extent, an attempt to classify neurons. The classification used morphological criteria (the only criteria available at the time) and buttressed its conclusions with remarkably modern cross-species comparisons. Neuronal classification remained a dominant theme in neurobiology over the following half-century but fell out of fashion as physiological and molecular methods matured, and mechanistic, 'hypothesis-driven' research came to be valued over projects with 'merely descriptive' aims. Over the past decade, however, many have come to believe that systematic categorization of cell types is an essential prerequisite for understanding mechanisms, and the descriptive enterprise has been revived[2–12]. Although we are proponents of this approach, we do not claim that such categorization will in and of itself lead to enlightenment. We do, however, believe that without it, enlightenment will be unattainable.

There have been two major obstacles to neuronal classification. The first is technical. Until recently, classification studies were both severely underpowered and highly biased owing to the laborious nature of the available methods. Compounding this issue, most methods were, at best, semi-quantitative, and problems of variance were crippling. In the past ten years, however, transformative advances have made it possible to analyse hundreds to tens of thousands of neurons quickly.

The second problem is conceptual. It is difficult to know how fine and firm the distinctions used to distinguish neuronal types from one another should be. It is indisputable that each neuron is different from every other neuron. If we take account of all the differences between neurons, however, the very notion of neuronal types becomes ephemeral. Conversely, although no one would doubt that broad classes of neurons are recognizable (motor neurons or cortical pyramidal cells, for example), these coarse distinctions are of little value for many experimental purposes. Is there a 'sweet spot' somewhere in between?

In this article, we tackle both of these issues. We describe recent attempts to circumvent the technical obstacles to cell-type classification using structural, functional and molecular criteria. To address the conceptual difficulties in the field, we use an analogy to another type of categorization — the classification of organisms into species — to suggest common-sense guidelines.

[1]*Allen Institute for Brain Science, Seattle, Washington, 98109, USA.*
[2]*Center for Brain Science and Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts, 02138, USA.*
sanesj@mcb.harvard.edu; hongkuiz@alleninstitute.org

doi:10.1038/nrn.2017.85
Published online 3 Aug 2017

Finally, we illustrate progress and remaining issues regarding neuronal classification using the retina and cerebral cortex as examples. These regions have been the subject of many classification studies, and lessons learned therein should be applicable to less fully investigated cell populations both in the brain and in other organs. Space limitations prevent us from discussing the classification of non-neuronal cells, but we note that many of the issues discussed here will be relevant for such efforts.

Purposes of cell-type classification

Few neuroscientists view neuronal classification as an end in itself. Rather, we hope that development of a cellular taxonomy will facilitate our understanding of how the brain works or, in diseases, fails to work properly. Designing a useful classification scheme therefore requires making explicit the needs it is meant to fulfil.

Reproducibility. One important purpose of neuronal classification is to enable the same types of neurons to be studied repeatedly. Long before molecular markers were available, investigators were able to map patterns of connectivity in invertebrates such as *Aplysia*, leech and *Caenorhabditis elegans*, successes that were enabled as much by the presence of uniquely identifiable neurons as by the circuits' simplicity. For vertebrates, with a few exceptions[13], neurons are not unique, and it is difficult to compare studies conducted in different places, at different times or with different methods. This problem was highlighted by Crick, who noted, "It is common for the experimentalist to record that, say, 25% of the neurons studied behave in a particular way, 37% in a different way and a further 15% in a third way. There is no indication…what type of neuron they are…. This is not science but rather natural history" (REF. 14). To answer this criticism, classification schemes must be sufficiently robust and easy enough to apply that they help settle arguments rather than cause new ones.

Genetic access. Increasingly, one goal of classification is to enable genetic access to specific cell types so that they can be marked and/or manipulated[15]. This is most straightforward for those species in which transgenesis is advanced – flies, worms, mice and zebrafish. However, new genome engineering methods and viral vectors are rapidly expanding the range of species to which genetic tools can be applied. Achieving this goal will require us to put a premium on molecular classification methods that lead to the discovery of genes that can be used as entry points.

Discovery. Classification leads directly to discoveries in several ways. As known cell types are catalogued, previously unrecognized ones may emerge. In addition, the profiling methods used for classification may lead to the identification of genes that become candidate determinants of cell type-specific morphology and function.

Understanding development. Elucidation of the steps by which a neuron acquires its form and function is difficult when it can be identified only after it has acquired those properties. As molecular markers have become available, immunohistochemical methods have transformed developmental neuroscience. Transgenic lines have improved the ability to prospectively identify neurons at early developmental stages. Molecular classification can identify new markers or marker combinations that drive this program forward. In addition, transcriptomic profiling of individual types can also provide candidate mediators of developmental choices.

Understanding evolution. As a fundamental unit of tissue and organismic function, cell types are invaluable for evolutionary comparisons[16]. For some cell types, conservation is obvious, and studies in one species are immediately applicable to others. In other cases, cell types that are unique to one of a pair of species may be responsible for critical differences between them. In contrast to our deep knowledge about the evolution and conservation of genes, our knowledge about the evolution and conservation of cell types is limited.

Studying disease. Some brain diseases primarily affect specific cell types. For example, amyotrophic lateral sclerosis affects upper and lower motor neurons[17], and congenital nystagmus affects starburst amacrine cells[18]. For others, the defects that lead to dysfunction remain mysterious. We might gain a better understanding of such disorders if we could find specific vulnerable neuronal types in disease models or autopsy materials. This type of inquiry will require high-throughput approaches, and for human tissue, only transcriptomic methods are likely to be applicable in the near future.

Generating a 'parts list'. Nowhere is the complexity of the brain more evident than in its enormous numbers of neurons and even greater numbers of synapses, both of which exhibit tremendous diversity. It seems obvious that we cannot make headway in addressing this complexity unless we consider neurons as types rather than as individuals — a crucial mode of dimensionality reduction. In addition, the shared functions of groups of neurons are likely to be a key organizing principle of brain function. It is therefore fair to say that we have no hope of understanding the brain without an accounting of cell types and their properties.

Defining neuronal types

In principle, it seems obvious that neurons should be viewed as members of a type if they serve a function that differs from the functions of other types of neurons. In practice, however, the functions of individual neurons can seldom be determined. Moreover, some functions may emerge only at the level of circuits. We therefore suggest that a more useful definition of type is a population of neurons with properties that are homogeneous within the population but differ from those of other neurons.

What are the relevant properties? The three main categories are morphological, physiological and molecular[2,5,6,9–11,19] (FIG. 1). Of the morphological properties of neurons, dendritic and axonal shapes and branching patterns have been the most informative; however, features

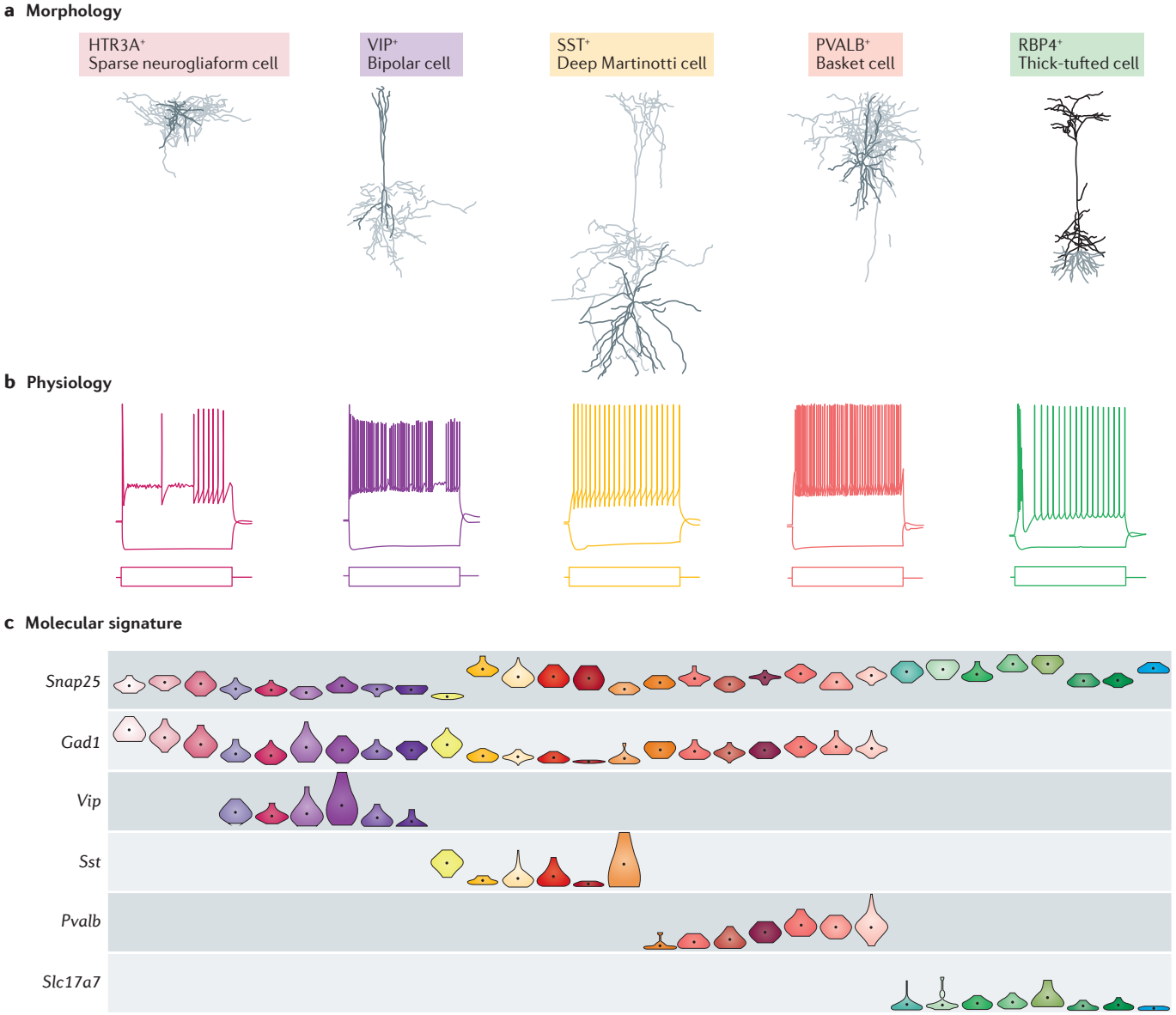


Figure 1 | **Criteria by which neurons can be classified.** Neurons can be classified using morphological, physiological and molecular criteria. **a** | Representative examples of five subclasses of cortical neurons obtained from brain slices. The cells were filled with biocytin, stained and imaged following patch clamp recording (see part **b**). Each subclass has distinct morphological features. For the four interneurons on the left, the dendrites are shown in dark grey and the axons in light grey. The soma of the 5-hydroxytryptamine receptor 3A-expressing (HTR3A$^+$) sparse neurogliaform cell is located in layer 1, and its axons are also concentrated in this layer. The vasoactive intestinal peptide-expressing (VIP$^+$) bipolar cell has a characteristic bipolar dendritic extension. The soma of the somatostatin-expressing (SST$^+$) deep Martinotti cell is located in layer 5/6, and its axons extend upward into layer 1. The parvalbumin-expressing (PVALB$^+$) basket cell has basket-like axonal arborisation. For the excitatory neuron on the right, the apical dendrites are shown in dark grey and the basal dendrites in light grey. This is a layer 5, thick-tufted cell from a retinol-binding protein 4 (*Rbp4*) gene promoter-driven Cre-expressing mouse. The cell features thick apical dendritic tufts extending into layer 1. These morphological features are consistent with those described in published reports[49,130,140]. **b** | Differential electrophysiological responses of the five subclasses of neurons shown in part **a** to square pulses of current in patch clamp recordings. For example, the HTR3A$^+$ cell is late spiking, whereas the PVALB$^+$ cell is fast spiking. These responses are consistent with those described in published reports[49,130,140]. **c** | Differential molecular signatures of the five subclasses of cortical neurons illustrated in part **a** derived from single-cell RNA-sequencing data. The violin plot shows the collective gene expression profile for each gene of all the cells in a type (cluster). We define the smallest discrete clusters of cells as types and the aggregates of types that share common features as classes or subclasses. Each transcriptomic cell type is shown as a column of data points with the same colour (the colour coding corresponds to that of the transcriptomic taxonomy shown in FIG. 5). Shown here are three interneuron cell types expressing *Htr3a* but not *Vip*, six interneuron cell types expressing *Vip*, six interneuron cell types expressing *Sst* and seven interneuron cell types expressing *Pvalb*. All of the interneurons express glutamate decarboxylase 1 (*Gad1*). Also shown are eight layer 5 excitatory neuron types, all of which express solute carrier family 17 member 7 (*Slc17a7*). All of the cells express synaptosome-associated protein 25 (*Snap25*). The height of each 'violin'-shaped data point represents the range of expression levels of the gene, and the width represents the proportion of cells displaying a particular level of expression. Parts **a** and **b** are from the Allen Cell Types Database (see Further Information). Part **c** is adapted with permission from REF. 136.

such as soma size and spine density are also used. The physiological properties of neurons include the resting potential, biophysical properties and the firing rate. Of the many molecular properties that can be considered, the most useful are protein composition (generally assayed immunohistochemically) and mRNA composition (measured by *in situ* hybridization and, increasingly, RNA sequencing). A fourth category of properties, connectivity, is equally relevant but is harder to assess and is therefore less often used.

In applying these criteria, we must consider the fact that no neuronal type is homogeneous. Heterogeneity arises from at least three sources. First, variations arise as cells develop and mature: genetic specification is incomplete, and cells that receive a single set of instructions (intrinsic and extrinsic) diversify as a result of a combination of stochastic and environmental differences[20,21]. Second, there are continuous, genetically encoded sources of variation, such as topographic gradients of 'mapping molecules' in the retina or tonotopic gradients of hair length in the cochlea[22,23]. Finally, variations result from ongoing changes in the adult environment, such as neural activity variations, hormonal fluctuations, or circadian rhythms[24,25]. These variations are genuine and meaningful; however, as we will argue below, they can often be distinguished from type-specific canonical properties.

Another problem is that a unifying definition of neuronal type should involve all three categories of properties, namely, physiological, morphological and molecular, which implies that they co-vary. Although such a satisfying correspondence has been shown for some types (see below), it seems unlikely that this will always be the case. When discrepancies between properties arise, we suggest that molecular criteria should be given interim precedence. This recommendation is mainly for practical reasons: as discussed above, for many purposes, molecular markers are currently the most useful. In addition, it is known that some molecular characteristics are enduring lifelong features of cells, whereas commonly measured physiological properties can vary depending on factors ranging from temperature to mood to sensory input. Cell non-autonomy also limits the utility of neuronal connectivity as a criterion: the loss of synaptic inputs or targets can alter a neuron's place in a circuit without causing any direct change to the neuron itself. Conversely, not all molecular properties are cell autonomous or static. For example, genes defining some stable properties (such as morphology) may be expressed only transiently during development, and the relevant molecular signatures may be undetectable in maturity. Nonetheless, it is reasonable to believe that there are permanent molecular features that maintain a cell's identity throughout an animal's life[4,26,27] and can thus serve as the basis for classification.

Neuronal types as species

In thinking about how to address the complexity of neuronal types, it may be useful to consult a field that groups individuals into types as its main preoccupation. In the field known as taxonomy, systematics or cladistics, the smallest discrete unit is the species. Although debates continue about how to define species and even whether they exist, systematics has nonetheless been a successful enterprise. The problems of defining species and neuronal cell types are similar in many ways[28–30], suggesting that there may be lessons to learn from the systematists.

There are three general schemes for defining species. The 'biological species concept' defines species boundaries in terms of reproductive isolation, which is not applicable to cell-type classification. Moreover, reproductive isolation in species is seldom tested experimentally and is therefore of far greater importance conceptually than practically, even within the field of systematics.

The second scheme for defining species is based on their phylogenetic relationships. Some attempts have been made to apply a similar method to neuronal cell types, substituting ontogeny for phylogeny. However, although it seems reasonable that related types of neurons would tend to descend from the same progenitors in a 'physical lineage tree', this is often not the case. In *C. elegans*, numerous sublineages generate motor neurons and a neuron's 'sibling' (that is, the other product of the terminal cell division that produced the neuron) is not particularly likely to be another cell of the same type[31]. In the vertebrate spinal cord, a motor neuron's sibling can be an astrocyte, an interneuron or another motor neuron[32]. In the retina, the same progenitor gives rise to all major classes of neuron; therefore, one type of ganglion cell is generally not a close relative of another ganglion cell type by lineage, although they are close relatives in terms of morphology, physiology and molecular architecture[33]. In short, the sorts of dendrograms of similarity that predict phylogenetic relationships for species do not predict lineage relationships for neuronal types.

An alternative instantiation of the phylogenetic parallel proposes a classification based on evolutionary conservation. For example, it has been proposed that a cell type should be defined as "a set of cells… that change in evolution together… and are evolutionarily more closely related to each other than to other cells" (REF. 16). In practice, however, there are few cases in which data are available to make this judgement. It may be more realistic to find ways to classify types within a species and then use that classification to launch an evolutionary inquiry.

Perhaps most useful is a third school of systematics, known as typological, taxonomic or phenetic systematics, which groups individuals into species according to their similarity of genotype and/or phenotype. There are several key principles that this approach advocates, which may provide lessons for cell-type classification[28–30]. First, group (that is, 'type') membership should be based on multiple criteria rather than on a single so-called 'essential' feature that the investigator favours. Second, the criteria for group membership should be rule-based, explicit and quantitative. Third, groupings should be hierarchical rather than flat to acknowledge the validity of both coarse and fine divisions (see below). Fourth, groupings generated by this approach should be viewed as hypotheses to be tested rather than

inflexible rules. Finally, classification should focus on discontinuities between groups and ignore parameters that vary continuously within what would otherwise be viewed as a group. Below, we will return to these concepts after considering the methods available to acquire the needed data.

High-throughput classification methods

Cell-type classification requires large data sets so that rare cell types can be found, and variation within cell types can be distinguished from differences between cell types. Past methods for comprehensive data collection were limited. In the past ten years, however, new methods have emerged that enable the collection of morphological, physiological and molecular data from large numbers of neurons (FIGS 2,3). Moreover, many of these methods are less biased than their predecessors, that is, they sample neurons in rough proportion to their actual frequency in the population.

Light microscopy. New methods for genetic sparse labelling and high-throughput light-microscopic imaging have fuelled efforts to assess neuronal morphologies, including entire axonal and dendritic arbours, in both fruitflies and mice. A key advance has been the generation of many cell type-specific transgenic lines driven by *cis*-regulatory elements in isolation or in a genomic context[34–43]. These and other tools are being used to collect whole-brain catalogues of morphologies[40,44–48] (FIG. 2a). In addition, recordings from and morphological reconstructions of thousands of neurons from cortical brain slices have been used to classify them into hundreds of morpho-electrical types[49]. The image data sets from both of these approaches can be used by specialized informatics pipelines to perform comparison and classification[50–52], although the low speed of reconstruction still poses a formidable bottleneck for neurons with intricate morphologies. Curation of these data in repositories, such as NeuroMorpho.Org (see Further Information), is also advancing[53].

Electron microscopy. Historically, electron microscopy was a powerful but laborious technique and the reconstruction of large tissue volumes was prohibitively time-consuming. Indeed, the 'nearly' complete connectome of *C. elegans*[54], reported in 1986, remained the only full reconstruction of more than a handful of cells for 25 years. However, advances in the field have systematically addressed each step in the process (FIG. 2b). Sectioning has been automated by block-face imaging[55,56] (in which the microtome used to section the tissue resides within the chamber of the microscope) and through the use of a tape-collecting microtome[57,58] (in which thousands of sections are collected automatically on a spool of tape). Imaging has been sped up by the use of multiple cameras that record separate parts of a large field simultaneously[59] and by the introduction of multi-beam scanning electron microscopes[60]. Reconstruction is benefiting from advances in machine learning that can segment images with less and less need for manual curation[61]. Conversely, the application of crowd-sourcing distributes the work among thousands of individuals[62,63]. Taking advantage of these innovations, reconstructions have been made to map the connectivity in the *Drosophila* optic and antennal lobes[64,65] and the mouse retina, thalamus and cortex[59,62,63,66–69].

Optical imaging of electrical activity. Classical electrode-based physiological methods collected data from one or, at most, a few dozen cells (using tetrodes) at a time. These methods are now being scaled up, allowing recording from hundreds of cells, with further increases foreseeable[70]. At present, however, optical imaging is the most effective approach for recording the activity of large numbers of neurons *in vivo* at the level of single cells (FIG. 2c). Molecular indicators of both voltage and calcium levels are available. In principle, voltage indicators are preferable because they measure neuronal activity more directly than calcium indicators and can detect subthreshold (synaptic) potentials[71]. In practice, however, only calcium indicators currently possess the sensitivity required for large-scale recordings *in vivo*[72]. Large-scale *in vivo* multi-photon calcium imaging has enabled the collection of functional information from hundreds to thousands of neurons simultaneously[73,74]. Imaging capability continues to grow rapidly to enable measurement from multiple areas[75,76] or areas deeper into the brain[77–81]. Compared with small and relatively transparent organisms (such as zebrafish[82,83]), multi-photon imaging is still restricted mostly to superficial structures (such as the cortex) of larger, mammalian brains. However, calcium imaging can be performed on cells 1 mm or deeper beneath the surface via gradient-index (GRIN) lenses or microendoscopes[84,85], which are beginning to be capable of resolving single cells.

Molecular profiling. A series of important advances in single-cell genome-wide molecular profiling techniques that have occurred over the past decade are benefiting cell-type classification efforts[25,86,87]. These advances include improvements in methods used to assess the genome, transcriptome, proteome and epigenome of single cells. All the indicated methods demonstrate vastly improved sensitivity compared with the previous generation of techniques. Single-cell profiling reveals diversity that is masked by averaging across populations, and genome-wide methods provide the unbiased coverage that was lacking in previous single-cell methods, such as quantitative PCR, immunohistochemistry or RNA *in situ* hybridization.

At this time, the most mature, scalable and useful technique for molecular profiling of cell type diversity is single-cell RNA sequencing (scRNA-seq) (FIG. 3a). This method is easily amenable to automation and, if applied at a sufficiently large scale, could drive a first 'complete' cell-type classification. To perform scRNA-seq, investigators dissociate cells from tissues and isolate them by manual picking, microfluidics or fluorescence-activated cell sorting. Subsequently, RNA is converted to cDNA, which is amplified and sequenced. A number of studies have used this approach to identify and classify cell types in a collection of neural tissues[25,88,89] (TABLE 1).

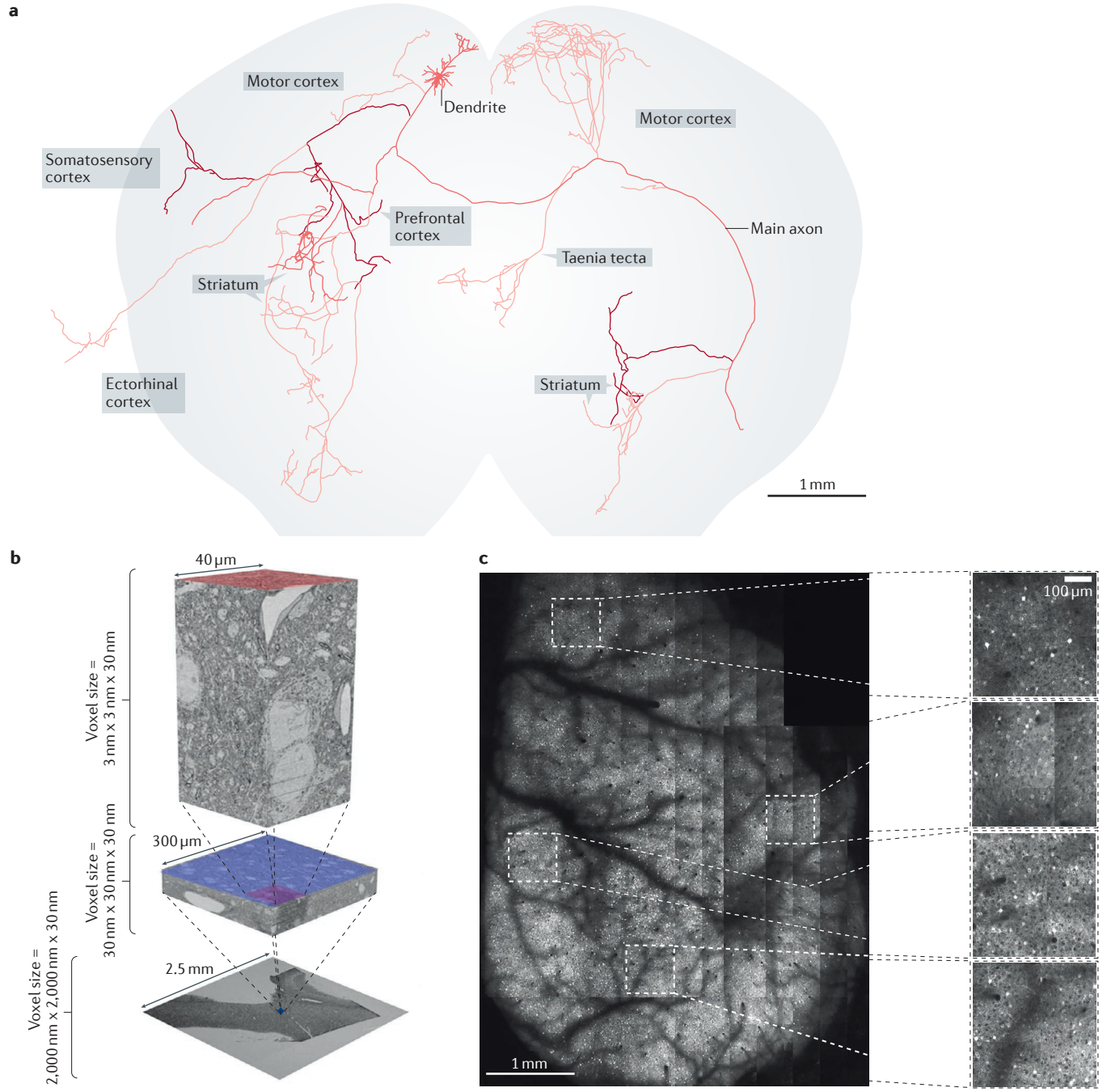


Figure 2 | **Imaging methods for profiling neuronal properties.** Large-scale imaging captures the structural and functional properties of populations of individual cells in great detail. **a** | High-throughput fluorescence light microscopy enables imaging of large tissue volumes (such as whole mouse brains) at high resolution, allowing visualization of complete neuronal morphologies. Shown here is a single intratelencephalic neuron reconstructed using the MouseLight system, a high-speed two-photon microscope that is integrated with a tissue vibratome[46]. The axons of this neuron project to multiple cortical and subcortical regions, as shown. Axonal arbours originating from common branch points are shown in the same shade of red. **b** | High-throughput electron microscopy combines methods for automated sectioning, imaging and reconstruction (segmentation) to reveal details of neuronal morphology and synaptic connectivity. The example shown illustrates multi-scale electron microscopy imaging of brain sections collected using the automated tape-collecting ultramicrotome[58]. The multi-scale imaging facilitates stitching and cross-registration and connects nanometre-range images with the larger tissue context. **c** | Large-scale fluorescence imaging methods enable the monitoring of the activity of hundreds or thousands of neurons through the use of organic dyes or genetically encoded indicators that report voltage or calcium levels. The image shows an example of results obtained using the Crystal Skull prep for long-term imaging[75]. In this preparation, a curved glass window replaces the mouse dorsal cranium, providing optical access to an estimated one million individual neurons across the dorsal surface of the neocortex for imaging of neuronal dynamics in behaving mice. The image is a mosaic of tiled two-photon images of a genetically encoded calcium indicator, GCaMP6f, in individual cells from the right hemisphere of a mouse. Inset images are magnified views of the corresponding boxed areas in the main panel. Part **a** is adapted with permission from REF. 46. Part **b** is adapted with permission from REF. 58. Part **c** is reproduced with permission from REF. 75.

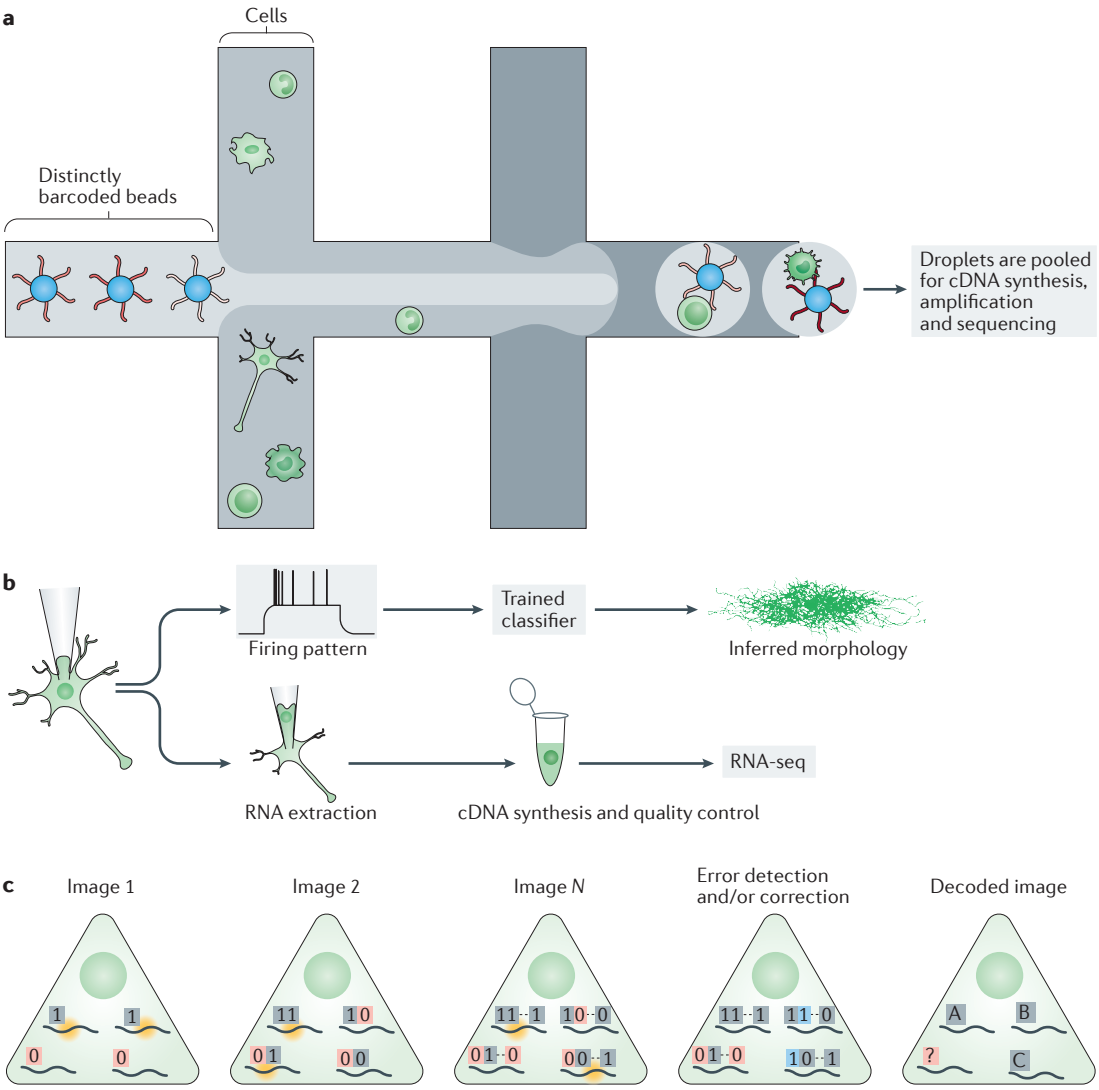


Figure 3 | **Molecular methods for profiling neuronal properties.** Single-cell genome-wide molecular profiling techniques provide unbiased and high-dimensional descriptions of molecular diversity. **a** | Large-scale single-cell RNA-sequencing (scRNA-seq) generates thousands of single-cell transcriptomic profiles. Shown here as an example is a schematic representing the Drop-seq method[97]. In this and related approaches (including InDrop[98] and Gemcode[99]), microfluidic devices pair individual cells with 'capture' reagents (beads that are 'barcoded' with oligonucleotides) that collect and barcode the cell's mRNAs in nanolitre-sized droplets. The droplets are then broken and reverse transcription, amplification and sequencing occurs in a pool of thousands of cells. **b** | Patch-seq methods extract mRNA from cells for scRNA-seq following electrophysiological recording, enabling direct correlation of molecular and physiological properties[107,108]. In addition to performing RNA-seq of extracted RNA, these methods compare the firing pattern of the cell to a 'trained classifier' (which summarizes the relationship between various firing patterns and morphological types) to infer the putative morphology of the recorded cell. **c** | Multiplexed fluorescence *in situ* hybridization (FISH) builds on single- and double-label methods to allow tens to hundreds of mRNAs to be co-assayed in individual cells within tissue sections. Multiplexed FISH can be applied to tissues following other types of profiling, allowing correlation of molecular profiles with morphological, physiological or functional properties. Shown here as an example is the multiplexed error-robust FISH (MERFISH) method[112], which uses combinatorial labelling and sequential imaging together with encoding schemes capable of detecting and/or correcting errors. Each RNA species is bound to oligonucleotide 'encoding' probes that label the RNA with a unique combination of N different 'readout sequences'. During each round of imaging, hybridization with a particular fluorescent 'readout' probe generates a signal (indicated by a yellow dot in the image) only in the subset of RNAs carrying the corresponding readout sequence, resulting in a binary code that reads 'on' (or '1') for probe binding and 'off' (or '0') for no binding. N rounds of imaging therefore generate a specific on/off (1/0) pattern for each RNA molecule, which is used to identify the RNA's localization and abundance. Such an encoding scheme allows highly multiplexed profiling of hundreds of genes. Importantly, an error correction method allows any error in detection (indicated by red shading) to be detected and, in many cases, corrected (indicated by blue shading).
Part **a** is adapted with permission from REF. 97. Part **b** is adapted with permission from REF. 107. Part **c** is from Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348, aaa6090 (2015). Reprinted/Adapted with permission from AAAS.

Table 1 | **Studies using single-cell RNA-sequencing to classify cell types in the nervous system**

Region or cell type	Number of cells sequenced	Number of cell types identified	Refs
Studies using single-cell tagged reverse transcription (Strt)			
Mouse DRG	799	11 (neuronal types)	152
Mouse S1 and CA1	3,005	47	153
Studies using Fluidigm C1 followed by Smart-seq			
Adult and fetal human cortex	466	10 classes and 7 neuronal types	154
Human cortex in development	393	NA (identified oRG-enriched genes)	155
Six human cortical areas	3,227	16	104
Juvenile mouse hypothalamus	898	62	156
Studies using Smart-seq			
Mouse DRG	203	10 types and 14 subtypes	157
Mouse V1	1,679	49	136
Studies using Smart-seq2			
Mouse midbrain LMX1A^{+} neural progenitors	550	NA (identified 2 lineages)	158
Studies using Drop-seq			
Mouse retina	44,808	39	97
Mouse retinal bipolar cells	~25,000	15	105
Mouse Arc-ME	20,921	50	159
Mouse EP	9,058	3 (neuronal types)	160
Human brain organoids	82,291	NA (identified diverse cortical and retinal cell types)	161

Arc-ME, hypothalamic arcuate-median eminence complex; DRG, dorsal root ganglion; EP, entopeduncular nucleus; NA, not applicable; oRG, outer radial glia; S1, primary somatosensory cortex; V1, primary visual cortex.

The methods currently in use fall into two main groups. In one group, the cells are dispersed in multi-well plates (one cell per well) and lysed. Poly(A) RNA is then reverse-transcribed to generate cDNA, which is amplified and used to generate a library that is eventually sequenced. The methods in this group include Smart-seq[90,91] (in which the full-length cDNA is fragmented for sequencing), single-cell tagged reverse transcription (Strt; in which only the 5′ end of the cDNA is barcoded and sequenced)[92,93] and cell expression by linear amplification and sequencing (CEL-seq; in which the cDNA is amplified linearly without PCR amplification, and only the 3′ end of the transcript is barcoded and sequenced)[94,95]. In a related approach, a microfluidic device called Fluidigm C1 sorts cells into micro-compartments, after which they are processed by Smart-seq or other protocols[96].

A second set of methods uses a microfluidic apparatus to pair single cells with oligonucleotide-bearing microspheres in nanolitre-sized aqueous droplets that are suspended in oil, generating an emulsion (FIG. 3a). Three current versions of the apparatus are known as Drop-seq[97], inDrop[98] and GemCode[99]. The cell is lysed within the droplet, and the mRNA is captured and reverse-transcribed. The emulsion is broken either before (Drop-seq) or after (inDrop and GemCode) reverse transcription, and amplification and library preparation occur in a single reaction for thousands of cells. The oligonucleotides are barcoded so that all mRNAs that arose from a single cell are indelibly marked and can be grouped after amplification. The advantage of this approach is the huge savings in cost and labour compared with generating libraries in thousands of individual wells. Conversely, the sequencing depth is generally low, ranging between 10K and 50K sequencing reads per cell for the droplet-based methods (although this is not a fixed limit) compared with millions of reads per cell for the plate-based methods.

Despite these differences, the methods share several features. For example, they all target poly(A)-tailed RNA species, thereby selectively capturing mRNAs and long non-coding RNAs. However, a new method, multiple annealing and dC-tailing-based quantitative single-cell RNA-seq (MATQ-seq), which is able to amplify all RNA species, has been developed[100]. Similarly, nearly all methods (except for Smart-seq) incorporate unique molecular identifiers that give each transcript a unique identity and thus allow for *in silico* correction of potential biases that may arise during PCR amplification. Conversely, Smart-seq and MATQ-seq preserve full-length transcript information and can therefore be used to identify alternative isoforms.

An exciting advance involves scRNA-seq from single nuclei[101–104]. Nuclei contain substantially less mRNA (mostly in the form of pre-mRNA) than somata, but the two compartments are similar with regards to gene representation; a main difference is that nuclei are biased towards recently transcribed genes and some specific types of mRNA. A main advantage of using nuclei is that they can be isolated from frozen or lightly fixed tissue by

gentle homogenization. This provides a means of obtaining single-cell data from tissues that cannot be freshly obtained (such as autopsy samples) or readily dissociated (such as heavily myelinated adult brain tissue).

In all these methods, a key question, given fixed resources, is whether to sequence more cells shallowly or fewer cells more deeply. One study[105] suggested that distributing a given number of reads over many cells may lead to a better resolution of cell types. However, obtaining large numbers of cells is sometimes infeasible. Thus, optimal methods depend on the situation and need: for example, shallow sequencing may be used for broad classification, and deeper sequencing used for sparser sampling or targeted populations. The task of systematically comparing scRNA-seq methods has begun[89,106], and it will be critical to conduct parallel studies to assess whether different methods arrive at the same cell types (also known as 'clusters'). Computational tools that can integrate data obtained with different methods are also needed.

Combining methods. Satisfactory cell-type classification requires the harmonization of morphological, physiological, molecular and possibly connectional categories. This is best accomplished by collecting two or more data types from the same cells. This was common in earlier generations of studies: for example, dye filling of neurons was performed following intracellular recording, and immunohistochemistry was performed on green fluorescent protein (GFP)-labelled cells. More recently, in Patch-seq (FIG. 3b), cellular contents are extracted following patch clamp recordings and subjected to scRNA-seq[107,108]. This combination provides rich data but is restricted to small numbers of cells. A related strategy is to 'spot-check' high-throughput data with a second method. Examples include imaging neurons by light or electron microscopy following calcium imaging[109] (physiology plus morphology) or *in situ* hybridization of sparsely labelled tissue based on scRNA-seq[105] (molecules plus morphology).

Scaling up to enable high-throughput analysis by multiple modalities presents greater challenges. Exciting developments in this area include multiplexed fluorescence *in situ* hybridization (FISH) and *in situ* sequencing methods that can be applied to tissue. The multiplexed FISH methods (FIG. 3c) that have been applied to brain sections include seqFISH[110,111], which was used to examine cell-type distribution patterns in the hippocampus, and multiplexed error-robust FISH (MERFISH), which can detect 140 genes at a low error rate and >1000 genes at a moderate error rate[112]. Newer versions of MERFISH incorporate modifications for high-throughput data generation[113] and background reduction[114]. Both seqFISH and MERFISH currently work on thin brain sections, whereas an expansion microscopy based FISH method, expansion FISH (ExFISH)[115], promises to detect gene expression in much thicker tissue blocks. *In situ* sequencing methods include fluorescent *in situ* sequencing (FISSEQ)[116] and padlock-probe-based rolling-circle amplification methods[117,118]. The multiplexed FISH and *in situ* sequencing approaches can simultaneously examine all the cells in a tissue sample for the expression of a set of preselected genes, acquiring essential information about the precise anatomical location and number and density of each investigated cell type that could not be obtained by scRNA-seq of cells from crudely dissected tissues.

Classification case studies

We next focus on the two parts of the mouse CNS to which these new high-throughput methods have been most intensively applied, the retina and cortex.

Retina. The neural retina contains five 'classes' of neurons, arranged in three layers, separated by two synaptic layers, as well as several classes of glia. The outer neuronal layer contains photoreceptors, which sense light. The middle layer contains three classes of interneurons, namely, horizontal cells, bipolar cells (BCs), and amacrine cells, which process the information and deliver it to retinal ganglion cells (RGCs) in the innermost layer. Axons of RGCs travel through the optic nerve, sending visual information to the brain. Studies have shown that each class of cells can be divided into multiple 'types'. The current estimate is that there are 100–150 retinal neuronal types[19,119,120].

Several features of the retina simplify the task of neuronal classification in this tissue. The retina has a clear laminar pattern, enabling neuronal classes to be identified based on position, and a relatively 'hard-wired' activity independent pattern of development. Moreover, the retina contains a complete circuit, enabling it to convert information from a precisely controllable sensory (visual) input to a single output with few (if any) retrograde connections. Of particular importance, most retinal neurons are arranged in a 'mosaic' pattern in which neurons of a single type are less closely spaced than would be expected by chance. Because neurons of a single type are randomly spaced relative to neurons of other types[121,122], statistical analysis of mosaic spacing provides a criterion for grouping neurons into types, independent of conventional structural, physiological and molecular properties. These features are not present in the cerebral cortex, accounting in part for the relative difficulty of categorizing cells in that tissue (see below).

Although cell-type categorization for all retinal neuronal classes is well underway, mouse BCs provide a particularly informative example of cells that have been successfully classified. BCs receive synapses from photoreceptors and horizontal cells on their dendrites and form synapses on RGCs and amacrine cells with their axons. Some receive inputs from rod photoreceptors and some from cone photoreceptors (rod and cone BCs, respectively). Cone BCs include cells that signal increases in light intensity and other cells that signal decreases in light; these cells are called ON and OFF BCs, respectively[120,123]. Three groups have classified BCs into types using some of the high-throughput methods described above: calcium imaging[124], crowd-sourced reconstruction of electron microscopy serial sections[62,63] and scRNA-seq by Drop-seq[105]. All the referenced studies used rigorous, quantitative criteria to group the cells into 14 or 15 types (FIG. 4a,b). Most importantly, light microscopic validation of the

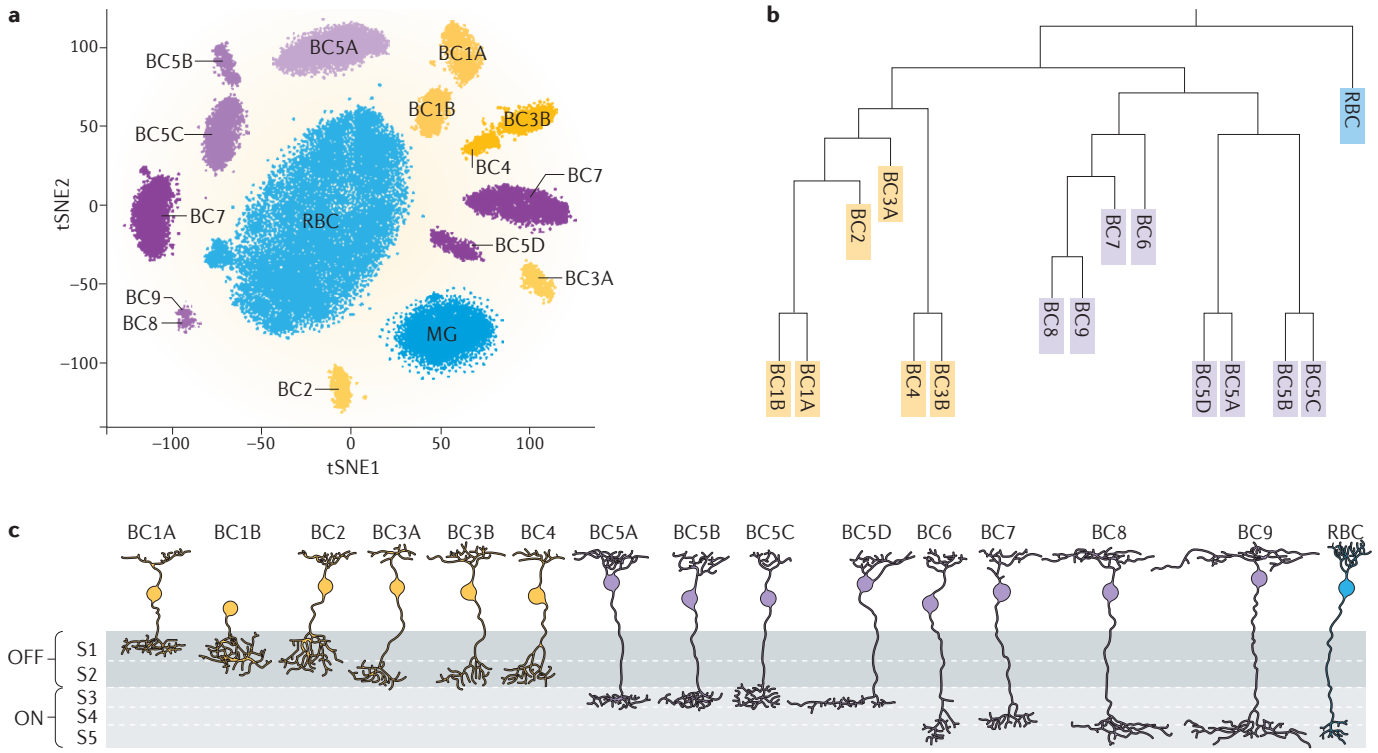


Figure 4 | **Classification of retinal bipolar cells.** Retinal bipolar cells (BCs) have been classified by three converging sets of high-throughput data, namely, morphological (electron microscopic reconstruction[62,63]), physiological (calcium imaging[124]) and molecular (Drop-seq[105]) data, into the following 15 types: one type of rod BC (RBC) and 14 types of cone BCs. The cone BCs are further subdivided into 8 ON types and 6 OFF types[105]. **a** | A t-distributed stochastic neighbour embedding (tSNE) plot showing clustering of ~20,000 BCs that were isolated by fluorescence-activated cell sorting from a visual system homeobox 2 gene promoter-driven green fluorescent protein (GFP)-expressing (*Vsx2-GFP*) transgenic mouse line, in which GFP is expressed in all BCs and Muller glial (MG) cells and profiled by Drop-seq. The tSNE plot provides a convenient way to display cell clusters, as defined by a high-dimensional analysis of correlations in gene expression, in two dimensions. **b** | Relationships among the BC types are shown in the form of a dendrogram that was created based on their transcriptomic similarity **c** | Sketches of the 15 BC types, whose terminal branches of axons are located in different sublaminae (S1-S5) of the inner plexiform layer in the retina. Parts **a–c** are adapted with permission from REF. 105.

physiologically and molecularly defined types provided strong evidence indicating that all the criteria converged on the same set of types (FIG. 4c). It therefore seems likely that all mouse BC types constituting more than ~1% of this class have now been identified.

These studies have yielded several insights. First, it was possible to assign virtually every cell to a single type, meaning that there was scant evidence for the existence of 'intermediate' forms of cells that could be assigned to two or more types equally well. Although this level of assignment may not generalize to other neuronal classes, it is comforting to know that discrete types can be identified through unbiased searches using multiple criteria. Second, the identified transcriptomic relationships paralleled the similarities that had been documented morphologically. For example, ON cone BCs are more closely related to each other than they are to OFF cone BCs, and vice versa, and cone BCs are more closely related to each other than they are to rod BCs (FIG. 4b). Third, the classification studies discovered new types. A previous authoritative study using morphological criteria and molecular markers described 11 BC types in mice[125]. In retrospect, three were missed because the available markers were limited; the authors suspected that additional types existed but were unable to demonstrate their existence. The final type, called 1B, was missed because it is actually unipolar (FIG. 4c) and was likely mistaken for an amacrine cell. However, this type has the molecular profile of a BC and lacks amacrine markers. Indeed, an independent study shows that its physiology and axonal ultrastructure are characteristic of BCs[126]. Moreover, with markers identified from scRNA-seq, it was possible to show that 1B cells are initially bipolar and then transform by late withdrawal of their dendritic processes.

Classification of other retinal classes is also proceeding, with discoveries as striking as those for BCs. For example, the number of mouse RGC types was estimated to be ~12 in 2004 (REF. 7); this number had increased to ~30 by 2015 (REF. 19) and has increased further to over 50 based on newer physiological studies[109], as well as preliminary data from ongoing electron microscopic reconstructions generated by crowdsourcing, such as those at http://museum.eyewire.org (Sebastian Seung, personal communication; see Further Information), and transcriptomic studies (J.R.S., unpublished observations). Some 20 types of amacrine cells have been characterized to date, and there is fragmentary evidence indicating that many more types exist (for example, 45 were reported in REF. 67).

Nonetheless, a complete classification of this region of the CNS is within sight, and based on the work discussed above, we predict that 100–150 types of cells will be identified.

Cerebral cortex. The cortex contains multiple sensory and motor areas as well as higher-order associational areas; the current number of different areas in humans, defined by cytoarchitecture and inferred connectivity (from neuroimaging), is estimated to be ~180 (REF. 127). Although all cortical areas have a laminar structure, the number and thickness of the layers vary across areas, and some cell types may be unique to specific areas. Most studies of neuronal types have focused on the rodent primary somatosensory and visual cortices.

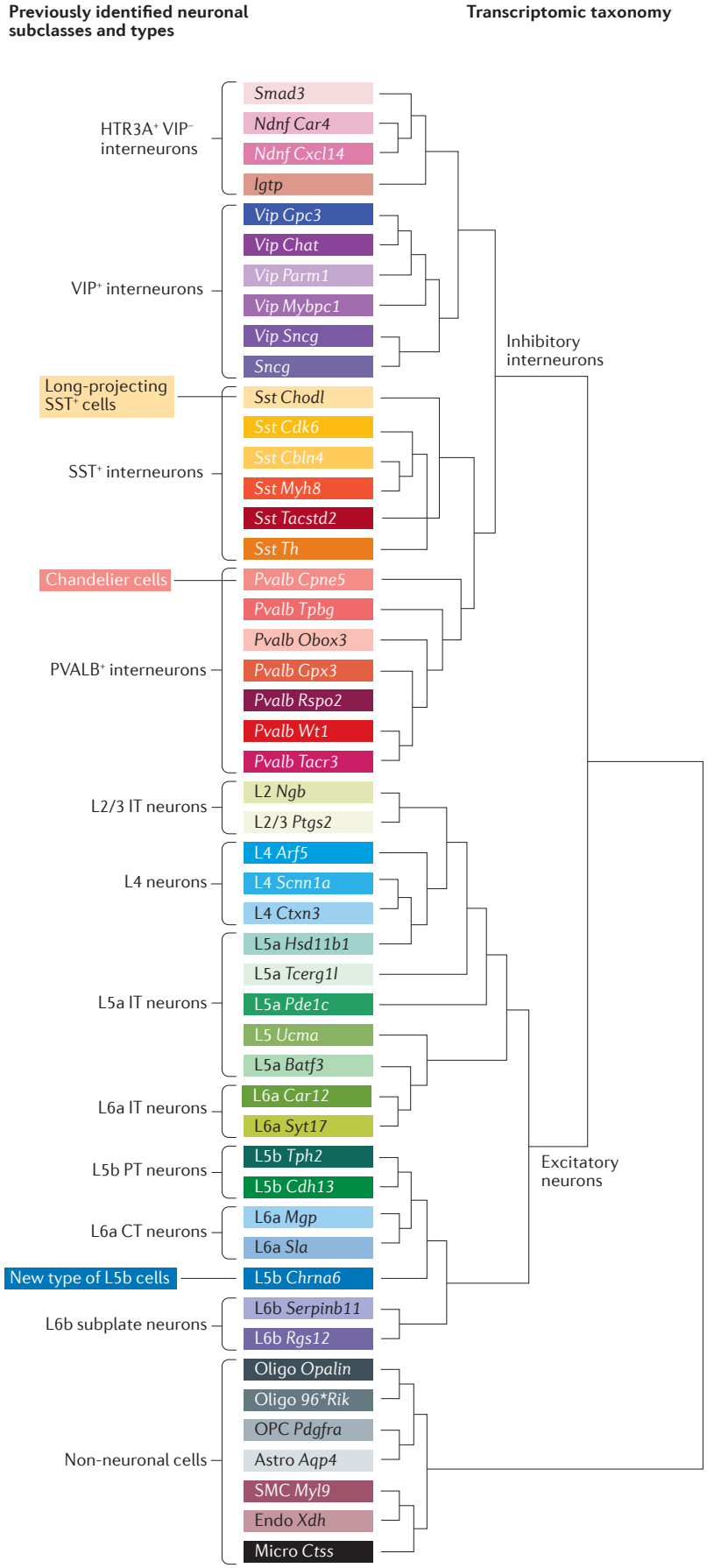
In general, cortical neurons are divided into two classes, namely, glutamatergic excitatory neurons and GABAergic inhibitory neurons, with multiple 'subclasses' within each class (FIGS 5,6). The following five excitatory neuron subclasses are known to exist: locally projecting layer 4 neurons, cortico-cortical projection neurons (also called callosal projection neurons or intratelencephalic neurons), subcerebral projection neurons (also called pyramidal tract neurons), cortico-thalamic projection neurons, and layer 6b subplate neurons[128–131]. The layer 4 neurons, which are the major postsynaptic targets of thalamic sensory nuclei, can be subdivided further into spiny stellate cells and star pyramidal cells, depending on whether they have an apical dendrite. Their axons and those of the layer 6b subplate neurons project locally or to nearby regions. The intratelencephalic neurons, which feature thin-tufted dendrites, are concentrated in layer 2/3 and upper layer 5 (often called layer 5a) but are also present in deeper layer 5 and layer 6. They project to multiple other cortical areas both ipsilaterally and contralaterally and to the striatum. The pyramidal tract neurons, which are located in relatively deeper layer 5 (layer 5b) and feature thick-tufted dendrites, project to multiple subcortical areas in the striatum, thalamus, midbrain, hindbrain and, sometimes, even down to the spinal cord. The cortico-thalamic neurons are located in layer 6, and they primarily provide feedback projections to the input thalamic nuclei. Many of the long-range projection neurons mentioned here also have local axonal collaterals that connect with other neurons in the same area, thus contributing to both local and global circuits.

It is unclear how many types of neurons constitute each of these five subclasses. The limited studies in which types have been defined by projection targets suggest that great variations exist with respect to projection specificity between different neurons[46,132–135], which is not surprising given the numerous combinations of targets from which an axon can choose. Conversely, the intrinsic electrophysiological properties of the excitatory cortical neurons exhibit relatively less variation than the properties of the other neurons. In a comprehensive morpho-electrical study of juvenile rat somatosensory cortical neurons, all excitatory neurons were assigned to a single 'e-type' (REF. 49). Single-cell transcriptomics has the potential to provide an exhaustive survey if enough cells are profiled. In fact, Smart-seq profiling of 1679 cells from adult mouse visual cortex revealed 19 excitatory neuronal types[136] (FIG. 5). It will be important to determine how these types are correlated with those determined by projection target specificity.

Within the GABAergic inhibitory class, there are four main subclasses, which are named for the neurochemical markers they express: parvalbumin-expressing ($PVALB^+$) cells, somatostatin-expressing (SST^+) cells, vasoactive intestinal peptide-expressing (VIP^+) cells and cells that express 5-hydroxytryptamine receptor 3A but lack VIP ($HTR3A^+VIP^-$)[137–140]. Each of these subclasses can be subdivided into types. For example, the $PVALB^+$ subclass can be split into basket cells and chandelier cells; the SST^+ subclass into Martinotti cells and non-Martinotti cells; the VIP^+ subclass into bipolar and multipolar cells; and the $HTR3A^+VIP^-$ subclass into neurogliaform cells and single bouquet cells. As indicated by their names, these interneuron types are defined by their axonal and dendritic morphologies, and each contains additional variations (similar to the hippocampus, wherein interneurons have been divided into at least 21 morphological types)[11,141]. In the past decade, the generation of many interneuron-specific GFP- or Cre recombinase-expressing transgenic mouse lines has dramatically increased our knowledge of the properties, connectivity and function of these interneuron types[140].

Nonetheless, it is still unclear exactly how many cell types constitute each inhibitory subclass. Studies in which cortical neurons were analysed physiologically (by patch clamp recording) and morphologically (by biocytin filling following recording) reported up to 194 interneuron types[49,142]. In the abovementioned scRNA-seq study of the adult mouse visual cortex, 23 interneuron types under the four main subclasses were identified and there was found to be some degree of correlation between these types and morpho-electrical types identified on the basis of known marker genes[136] (FIG. 5).

Our current understanding of the cell types in the cortex is similar to our understanding of the cell types in the retina in some ways but differs in others. Similar to studies of the retina, preliminary comprehensive scRNA-seq studies (with more neurons and more cortical areas than previously published)[136] of the cortex suggest that the number of cell types found therein is currently in the 100–200 range (H.Z., unpublished observations). However, in the cortex, no clear patterns of arrangement corresponding to the mosaicism or tiling seen in the retina have been identified. Cells intermediate between transcriptomic types[136], as well as orthogonality between cell types defined by electrophysiology and those defined morphologically[49], have been observed in cortex. Further studies will be needed to determine to what extent their prevalence varies between different regions. It is hoped that the arsenal of new high-throughput methods mentioned above will eventually make it possible to reach a census regarding the classification of all cortical cell types, in much the same way as is being done for retina. As explained below, we expect that a hierarchical taxonomy will be the most appropriate way to encapsulate the various degrees of distinction and relatedness between cortical cell types and to assess the relationships between cortical areas.



Principles of cell-type classification

Above, we noted the parallels between the classification of individual organisms into species and the classification of neurons into types. Studies on the retina, cortex and other regions suggest that neuronal classification would benefit from the application of taxonomic principles in the following three specific ways:

Figure 5 | **Neuronal classes and types found in the cerebral cortex.** Cell types defined by single-cell transcriptomics (centre) are annotated with neuronal subclasses and types previously identified using morphological, connectional and neurochemical properties (left). The transcriptomic taxonomy 'tree' (right) was established by iterative clustering analysis of the single-cell RNA-sequencing data from 1679 primary visual cortical cells[136]. Each coloured box in the transcriptomic taxonomy represents one transcriptomic type, named after the typical marker genes shown in the box. This study identified 49 transcriptomic types, including 23 inhibitory interneuron types, 19 excitatory neuron types and 7 non-neuronal types. The neuronal types correspond to four inhibitory and seven excitatory subclasses that were previously known. The boxed labels on the left show examples of special or new cell types that may correspond to specific transcriptomic types based on their branch location on the taxonomy tree and thus reveal new marker genes for these unique neuronal types. For example, somatostatin-expressing (SST+) cells with long-range axonal projections may correspond to the *Sst* chondrolectin (*Chodl*) type, chandelier cells may correspond to the parvalbumin (*Pvalb*) copine 5 (*Cpne5*) type, and the layer 5b (L5b) cholinergic receptor nicotinic alpha 6 subunit (*Chrna6*) type may represent a new type of excitatory neuron. 96*Rik, 9630013A20Rik; *Aqp4*, aquaporin 4; *Arf5*, ADP ribosylation factor 5; Astro, astrocyte; *Car4*, carbonic anhydrase 4; *Cbln4*, cerebellin 4 precursor; *Cdh13*, cadherin 13; *Cdk6*, cyclin-dependent kinase 6; *Chat*, choline O-acetyltransferase; CT, cortico-thalamic; *Ctss*, cathepsin S; *Ctxn3*, cortexin-3; *Cxcl14*, C-X-C motif chemokine ligand 14; Endo, endothelial cell; *Gpc3*, glypican-3; *Gpx3*, glutathione peroxidase 3; *Hsd11b1*, hydroxysteroid 11-beta dehydrogenase 1; *Igtp*, interferon gamma-induced GTPase; IT, intratelencephalic; *Mgp*, matrix Gla protein; Micro, microglia; *Mybpc1*, myosin-binding protein C, slow type; *Myh8*, myosin heavy chain 8; *Myl9*, myosin light chain 9; *Ngb*, neuroglobin; Oligo, oligodendrocyte; OPC, oligodendrocyte progenitor cell; *Parm1*, prostate androgen-regulated mucin-like protein 1; *Pde1c*, phosphodiesterase 1c; *Pdgfra*, platelet derived growth factor receptor alpha; PT, pyramidal tract; *Ptgs2*, prostaglandin-endoperoxide synthase 2; *Rgs12*, regulator of G protein signalling 12; *Rspo2*, R-spondin 2; *Scnn1a*, sodium channel epithelial 1 alpha subunit; *Sla*, Src like adaptor; SMC, smooth muscle cell; *Sncg*, synuclein gamma; *Syt17*, synaptotagmin 17; *Tacr3*, tachykinin receptor 3; *Tacstd2*, tumour-associated calcium signal transducer 2; *Tcerg1l*, transcription elongation regulator 1 like; *Th*, tyrosine hydroxylase; *Tpbg*, trophoblast glycoprotein; *Tph2*, tryptophan hydroxylase 2; *Ucma*, upper zone of growth plate and cartilage matrix associated; *Vip*, vasoactive intestinal peptide; *Wt1*, Wilms tumour 1; *Xdh*, xanthine dehydrogenase. Figure adapted with permission from REF. 136.

Criteria. Taxonomists stress that group membership should be based on multiple criteria rather than a single 'essential' feature. A corollary is that groups should not be named for the supposedly essential feature, as this will often end up being misleading. Indeed, there may be no single characteristic that is possessed by every group member. As noted by Tyner, a pioneer in applying insights from systematics to neuroscience, "A biological population can be described in terms of a substantial list of features such that (a) each member of the group possesses a large… number of the features; (b) each feature is possessed by a large… number of individuals in the group; and (c) no feature is necessarily possessed by every member of the group" (REF. 30).

It is also important that the criteria for defining types be quantitative. Historically, qualitative descriptions were deemed to suffice. These included morphological descriptions based on Golgi staining, physiological divisions based on slowly and rapidly adapting electrophysiological responses and molecular divisions into cells that express high or low levels of a marker. Fortunately, many newer high-throughput methods are intrinsically quantitative, facilitating the use of robust clustering algorithms to place individuals into groups. This switch may dispel some of the scepticism about the feasibility of drawing clear distinctions between types.

Continuous and discontinuous variations. Variation in most quantifiable features used to classify organisms or neurons is inevitable. As noted above, continuous variation can arise from genetic, environmental or stochastic sources. We suggest that continuous and discontinuous variation should be treated differently and that the latter is decisive in defining neuronal types.

This point can be illustrated at least conceptually by considering the t-distributed stochastic neighbour embedding (tSNE) plot, which demonstrates the results of a transcriptomic analysis of retinal BCs (FIG. 4a). In this type of plot, each dot represents one cell, and the distance between two cells is related to their gene-expression similarity in a multidimensional, nonlinear space. There is considerable heterogeneity with respect to gene expression between the cells in each cluster, and this heterogeneity arises from both biological and technical factors; however, no studies to date have been able to further subdivide even the most numerous type, namely, rod BCs. The important point is that the extensive, continuous transcriptional variation among rod BCs does not lead to further subdivision of its cluster or preclude clear separation from neighbouring clusters, which comprise cone BC types. One cannot interpret distances within a cluster in an intuitive way because they are influenced by cell number. The lesson, however, is that focusing on discontinuous variations results in meaningful categorization, whereas focusing on continuous variation is more problematic.

Another example is the direction and/or orientation selectivity that is exhibited by RGCs and visual cortical neurons. Many visual cortical neurons selectively respond to a specific orientation of visual stimuli; at the population level, such selectivity varies continuously between individual cells such that, collectively, they represent all directions[143]. By contrast, each retinal direction-selective RGC responds to one of four discrete directions: dorsal, ventral, nasal or temporal[144,145]. As such, according to our working definition, the preferred orientation or direction would be a criterion for classifying RGC types but is not useful for classifying cortical neuron types. Indeed, this distinction corresponds well with RGC or cortical neuron types defined by other (molecular or morphological) parameters[19,143,146].

An ancillary benefit of stressing discontinuous variation is that it may provide a way to take into account many (although certainly not all) of the 'state' properties of neurons – properties altered by activity, hormonal milieu, circadian rhythms and a host of other factors. Many of these properties vary continuously and can thus be distinguished from the more permanent, canonical properties that can be used to define types. Although some features also vary continuously in a way that spans multiple types and may be useful for classification[109], they currently remain a second choice for classification if discontinuous variables are available.

Hierarchical classification. Use of hierarchical (multi-level) rather than flat (single-level) classification systems has the following two advantages: it includes relationships between types as an intrinsic feature of the classification, and it provides a flexible way to update the system in light of new information. FIGURE 6 shows the hierarchical schemes that we propose for the retina and cortex, and they are consistent with previous definitions. We suggest that groups should be defined within specific anatomical brain regions. The smallest discrete groups (which notionally serve a specific function) should be called types, and the largest aggregates of types that share common functional features should be called classes. In between types and classes are subclasses, which can have one or more levels. According to this scheme, a cell can be assigned to multiple groups but with only one group at each level. The function of many of these cell types is elusive, but the retina provides some instructive examples in which a type can be correlated with a particular function. For instance, the BC class is tasked with the broad function of collecting input from photoreceptors and delivering it to RGCs, whereas each individual BC type has the more specific function of carrying input from rods or cones and converting light signals into sustained or transient excitation or inhibition.

In the future, as additional classes are defined and compared, higher levels may emerge, for example, sensory, interneuron and projection classes, as well as groupings that span multiple regions. Finer distinctions within types can be accommodated as subtypes, just as species are now divided into strains. Subtypes may also provide a useful way to categorize cells in which some properties are fixed while others vary between distinct states, for example, owing to transcriptional alterations[147] or neurotransmitter switching[148]. In many cases, we expect that subtypes will prove to be provisional categories: in some cases, enough distinctions will emerge for these to be viewed as authentic types, whereas in other cases, as larger numbers of cells are profiled, subtypes will merge.

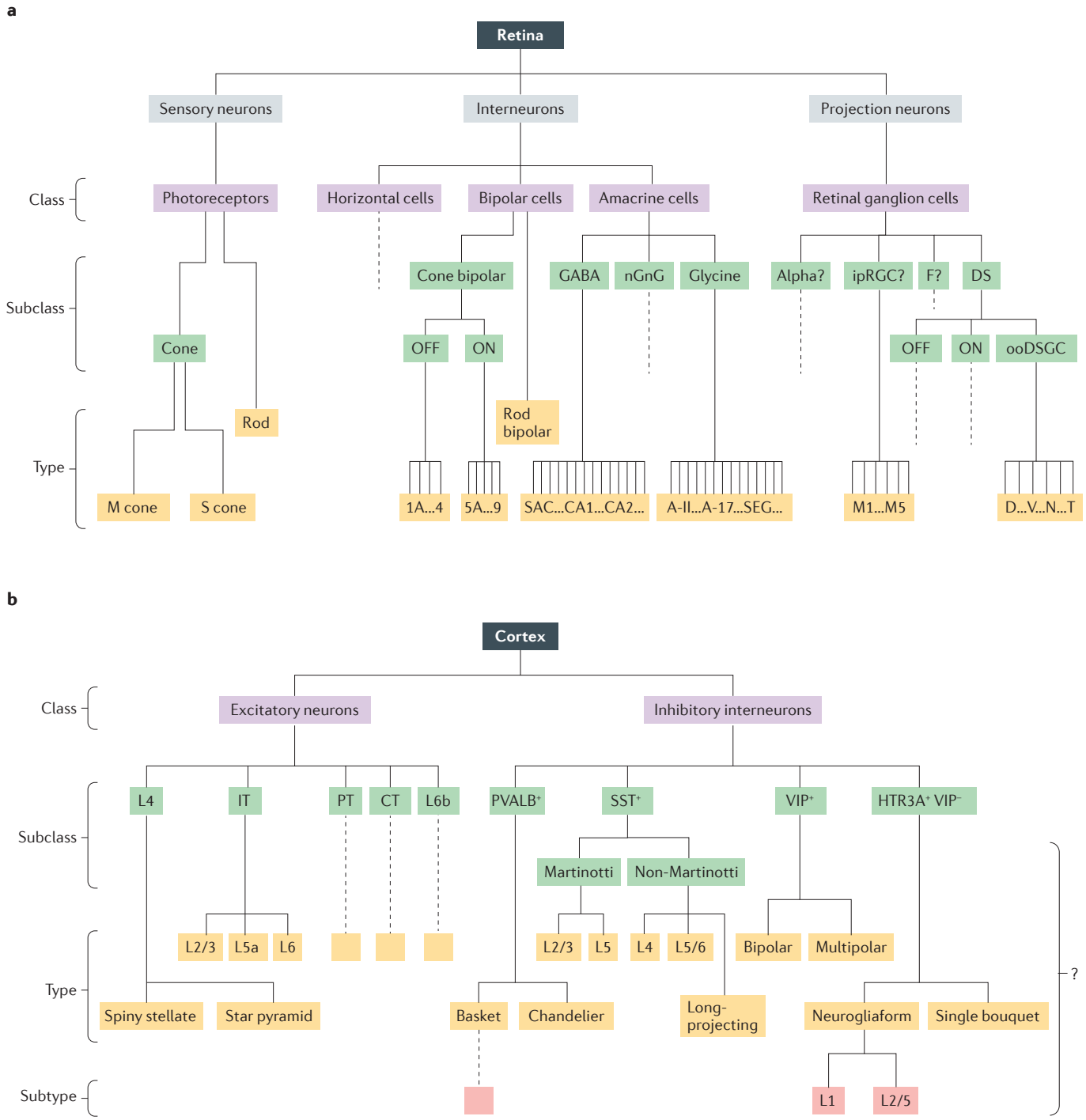


Figure 6 | **Hierarchical classification of neurons.** The figure shows a proposed hierarchical classification of cells in the retina (**a**) and cerebral cortex (**b**). In both areas, individual cell types can be grouped into classes, and intermediate levels of subclasses can be determined based on distinct morphological, physiological and molecular features. Higher-order groupings (such as those shown in part **a**, including sensory neurons, interneurons and projection neurons) may emerge once enough areas have been provided and compared. Types are the commonly recognized ('validated') terminal branches in the current hierarchical arrangement of cell types. Lower-order groupings into subtypes may largely be provisional until additional data are collected that could determine if they could form new types or should be merged into other types. Dashed lines indicate the presence of additional types that cannot be labelled due to lack of space. The question marks in part **a** indicate that the hierarchical relationship among the indicated cell types remains unclear. The question mark in part **b** indicates that the status of the cortical cell groups indicated may be either subclasses, types or subtypes. CT, cortico-thalamic neurons; DS, direction-selective retinal ganglion cells (RGCs); F, forkhead box P2 (Foxp2)-expressing RGCs; HTR3A, 5-hydroxytryptamine receptor 3A; ipRGC, intrinsically photosensitive RGCs; IT, intratelencephalic neurons; L4, layer 4; L6b, layer 6b subplate neurons; nGnG, non-GABAergic-non-glycinergic amacrine cells; ooDSGC, ON-OFF direction-selective RGCs; PT, pyramidal tract neurons; PVALB, parvalbumin; SST, somatostatin; VIP, vasoactive intestinal peptide.

A path forward

Classification requires us to describe and establish correspondence between molecular composition, morphology, connectivity, physiology and function. It is clear, however, that these properties are too complex and variable to address them all at once. From a practical point of view, we envision classifying neuronal cells into five overlapping stages, which can be completed within the next decade.

The first stage will involve quantitatively classifying neurons in individual regions of the nervous system by molecular and morphological criteria and attempting to establish correspondence between the two. The retinal and cortical studies described above will serve as a model for these endeavours. Molecular classification will rely on increasingly affordable high-throughput single-cell transcriptomics. Likewise, high-throughput technologies in light and electron microscopy can be deployed to generate large-scale data sets for the latter. Correlations between morphology and molecular identity may be best established by multiplexed FISH or *in situ* sequencing. Such correlations, if established, would facilitate the attainment of a baseline understanding of neuronal types.

In the second stage, connectivity (inputs and outputs) and physiology will be incorporated into the classification scheme. Current technologies are limited in their ability to collect comprehensive connectional and physiological data for large populations of neurons, but calcium imaging and virus-mediated tracing methods are improving rapidly. Continued development of these and other technologies and their unbiased application will enable systematic connectional, physiological and functional characterization of individual cells. Ultimately, this will be informative and useful not only in refining the cell-type classification but also in understanding the relationship between a cell's variable states and its core cell-type identity.

Third, as classifications within regions become authoritative, it will be possible to make comparisons across areas. This will enable researchers to determine, for example, whether the cell types in the auditory cortex are slight variants of those in visual cortex or whether there are fundamental differences between the two cell types. Likewise, researchers will be able to determine whether or not the excitatory neurons in the cortical and subcortical regions are close relatives. These comparisons will also be useful in designing intersectional strategies to access specific types in specific regions. Furthermore, a complete cell-type inventory needs to include a description of the types, proportions and spatial distributions of cells in different regions. Such information will help us understand the larger-scale organization of the nervous system.

Fourth, classification schemes derived from healthy adult animals will be used as a foundation for investigating development, evolution and disease. These comparisons will help us understand how cell types emerge in development and the extent to which they are conserved phylogenetically. We will be able to answer questions about how immature neurons diversify to acquire their adult fates and how cell-type identity is determined and maintained. We may be able to determine which neuronal types are evolutionarily conserved and which are unique to individual families or species. Comparisons can be extended further to diseases and disease models to understand to what extent disease-related alterations are confined to specific neuronal types and how animal models are related to human conditions.

Finally, it will be important to develop a unified nomenclature for neural cell types. Optimally, this nomenclature should be applicable across brain regions and species and should incorporate molecular, morphological, physiological and, perhaps, connectional criteria. If this nomenclature is to be useful, it will need to be widely adopted by the community. These are all difficult challenges, and it may be several years until enough data are available to formulate a proposal. In the meantime, we make some modest suggestions. Classification should be based on a hierarchical scheme (as described above) so that additional groupings can be added or interpolated as new information becomes available. The names of classes and types should avoid reference to putative functions to facilitate integration across regions and species. Finally, any proposal should be made by a group of neuroscientists whose expertise spans multiple regions, species and technologies. Models are provided by committees that propose nomenclature for genes, enzymes and receptors[149–151]. Their success encourages us to believe that a similar mechanism could eventually be employed to design a cell-type nomenclature that would be both generally useful and widely adopted.

Overall, the enterprise of cell-type classification has the potential to transform our view of nervous system function and malfunction. Many problems remain, and, as is the case for species taxonomy, cell-type classification schemes need to be regarded as hypotheses to be tested and refined as we move through these stages. Nonetheless, the pace of progress is rapid, and some of the conceptual and technical challenges that seem formidable today are likely to be surmounted over the next decade.

1. Ramón y Cajal, S. *Histologie Du Système Nerveux de L'homme & Des Vertébrés* (ed. Maloine, A.) (Paris: Maloine, 1909). Translated by Swanson, N. & Swanson, L. W. (Oxford University Press, 1995). **The founding document of neurobiology, which was in large part an attempt to classify neuronal types with the best method available at the time, Golgi staining.**
2. Ascoli, G. A. *et al.* Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat. Rev. Neurosci.* **9**, 557–568 (2008).
3. Bota, M. & Swanson, L. W. The neuron classification problem. *Brain Res. Rev.* **56**, 79–88 (2007).
4. Brenner, S. Sequences and consequences. *Phil. Trans. R. Soc. B* **365**, 207–212 (2010).
5. Fishell, G. & Heintz, N. The neuron identity problem: form meets function. *Neuron* **80**, 602–612 (2013).
6. Markram, H. *et al.* Interneurons of the neocortical inhibitory system. *Nat. Rev. Neurosci.* **5**, 793–807 (2004).
7. Masland, R. H. Neuronal cell types. *Curr. Biol.* **14**, R497–R500 (2004).
8. Migliore, M. & Shepherd, G. M. An integrated approach to classifying neuronal phenotypes. *Nat. Rev. Neurosci.* **6**, 810–818 (2005).
9. Nelson, S. B., Sugino, K. & Hempel, C. M. The problem of neuronal cell types: a physiological genomics approach. *Trends Neurosci.* **29**, 339–345 (2006).
10. Seung, H. S. & Sumbul, U. Neuronal cell types and connectivity: lessons from the retina. *Neuron* **83**, 1262–1272 (2014).
11. Somogyi, P. & Klausberger, T. Defined types of cortical interneurone structure space and spike timing in the hippocampus. *J. Physiol.* **562**, 9–26 (2005).

12. Stevens, C. F. Neuronal diversity: too many cell types for comfort? *Curr. Biol.* **8**, R708–R710 (1998).
13. Korn, H. & Faber, D. S. The Mauthner cell half a century later: a neurobiological model for decision-making? *Neuron* **47**, 13–28 (2005).
14. Crick, F. The impact of molecular biology on neuroscience. *Phil. Trans. R. Soc. Lond. B* **354**, 2021–2025 (1999).
15. Huang, Z. J. & Zeng, H. Genetic approaches to neural circuits in the mouse. *Annu. Rev. Neurosci.* **36**, 183–215 (2013).
This article summarizes the progress enabled by obtaining genetic access to specific cell types, which relied on molecular classification.
16. Arendt, D. *et al.* The origin and evolution of cell types. *Nat. Rev. Genet.* **17**, 744–757 (2016).
17. Taylor, J. P., Brown, R. H. Jr & Cleveland, D. W. Decoding ALS: from genes to mechanism. *Nature* **539**, 197–206 (2016).
18. Yonehara, K. *et al.* Congenital nystagmus gene FRMD7 is necessary for establishing a neuronal circuit asymmetry for direction selectivity. *Neuron* **89**, 177–193 (2016).
19. Sanes, J. R. & Masland, R. H. The types of retinal ganglion cells: current status and implications for neuronal classification. *Annu. Rev. Neurosci.* **38**, 221–246 (2015).
This article reviews the progress in classifying RGC types in mice. Remarkably, the estimated number has grown from ~30 to >50 in the two years since its publication.
20. Anderson, D. J. The neural crest cell lineage problem: neuropoiesis? *Neuron* **3**, 1–12 (1989).
21. Boije, H., Rulands, S., Dudczig, S., Simons, B. D. & Harris, W. A. The independent probabilistic firing of transcription factors: a paradigm for clonal variability in the zebrafish retina. *Dev. Cell* **34**, 532–543 (2015).
22. McLaughlin, T. & O'Leary, D. D. Molecular gradients and development of retinotopic maps. *Annu. Rev. Neurosci.* **28**, 327–355 (2005).
23. Zipursky, S. L. & Sanes, J. R. Chemoaffinity revisited: dscams, protocadherins, and neural circuit assembly. *Cell* **143**, 343–353 (2010).
24. Rowe, M. H. & Stone, J. The interpretation of variation in the classification of nerve cells. *Brain Behav. Evol.* **17**, 123–151 (1980).
25. Wagner, A., Regev, A. & Yosef, N. Revealing the vectors of cellular identity with single-cell genomics. *Nat. Biotechnol.* **34**, 1145–1160 (2016).
26. Deneris, E. S. & Hobert, O. Maintenance of postmitotic neuronal cell identity. *Nat. Neurosci.* **17**, 899–907 (2014).
27. Waddington, C. H. *The strategy of the genes: a discussion of some aspects of theoretical biology* (Allen & Unwin, 1957).
28. Rodieck, R. W. & Brening, R. K. Retinal ganglion cells: properties, types, genera, pathways and trans-species comparisons. *Brain Behav. Evol.* **23**, 121–164 (1983).
29. Rowe, M. H. & Stone, J. Naming of neurones. Classification and naming of cat retinal ganglion cells. *Brain Behav. Evol.* **14**, 185–216 (1977).
30. Tyner, C. F. The naming of neurons: applications of taxonomic theory to the study of cellular populations. *Brain Behav. Evol.* **12**, 75–96 (1975).
A prescient discussion of how taxonomic principles could be used to improve cell-type classification. Cast in terms of physiological properties (see also REFS 24, 28 and 29), it is even more applicable to modern molecular schemes.
31. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode *Caenorhabditis elegans. Dev. Biol.* **100**, 64–119 (1983).
32. Leber, S. M., Breedlove, S. M. & Sanes, J. R. Lineage, arrangement, and death of clonally related motoneurons in chick spinal cord. *J. Neurosci.* **10**, 2451–2462 (1990).
33. Agathocleous, M. & Harris, W. A. From progenitors to differentiated cells in the vertebrate retina. *Annu. Rev. Cell Dev. Biol.* **25**, 45–69 (2009).
34. Ma, Y., Hu, H., Berrebi, A. S., Mathers, P. H. & Agmon, A. Distinct subtypes of somatostatin-containing neocortical interneurons revealed in transgenic mice. *J. Neurosci.* **26**, 5069–5082 (2006).
35. Dietzl, G. *et al.* A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* **448**, 151–156 (2007).
36. Gong, S. *et al.* Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. *J. Neurosci.* **27**, 9817–9823 (2007).
37. Kim, I. J., Zhang, Y., Yamagata, M., Meister, M. & Sanes, J. R. Molecular identification of a retinal cell type that responds to upward motion. *Nature* **452**, 478–482 (2008).
38. Siegert, S. *et al.* Genetic address book for retinal cell types. *Nat. Neurosci.* **12**, 1197–1204 (2009).
39. Taniguchi, H. *et al.* A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* **71**, 995–1013 (2011).
40. Jenett, A. *et al.* A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* **2**, 991–1001 (2012).
41. Manning, L. *et al.* A resource for manipulating gene expression and analyzing *cis*-regulatory modules in the *Drosophila* CNS. *Cell Rep.* **2**, 1002–1013 (2012).
42. Gerfen, C. R., Paletzki, R. & Heintz, N. GENSAT BAC cre-recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits. *Neuron* **80**, 1368–1383 (2013).
43. Harris, J. A. *et al.* Anatomical characterization of Cre driver mice for neural circuit mapping and manipulation. *Front. Neural Circuits* **8**, 76 (2014).
44. Chiang, A. S. *et al.* Three-dimensional reconstruction of brain-wide wiring networks in *Drosophila* at single-cell resolution. *Curr. Biol.* **21**, 1–11 (2011).
45. Shih, C. T. *et al.* Connectomics-based analysis of information flow in the *Drosophila* brain. *Curr. Biol.* **25**, 1249–1258 (2015).
46. Economo, M. N. *et al.* A platform for brain-wide imaging and reconstruction of individual neurons. *eLife* **5**, e10566 (2016).
47. Gong, H. *et al.* High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. *Nat. Commun.* **7**, 12142 (2016).
48. Gong, H. *et al.* Continuously tracing brain-wide long-distance axonal projections in mice at a one-micron voxel resolution. *Neuroimage* **74**, 87–98 (2013).
49. Markram, H. *et al.* Reconstruction and simulation of neocortical microcircuitry. *Cell* **163**, 456–492 (2015).
This paper reports a large-scale electrophysiological and morphological characterization and classification of neocortical neurons.
50. Costa, M., Manton, J. D., Ostrovsky, A. D., Prohaska, S. & Jefferis, G. S. NBLAST: rapid, sensitive comparison of neuronal structure and construction of neuron family databases. *Neuron* **91**, 293–311 (2016).
51. Peng, H. *et al.* BigNeuron: large-scale 3D neuron reconstruction from optical microscopy images. *Neuron* **87**, 252–256 (2015).
52. Sumbul, U. *et al.* A genetic and computational approach to structurally classify neuronal types. *Nat. Commun.* **5**, 3512 (2014).
53. Ascoli, G. A., Maraver, P., Nanda, S., Polavaram, S. & Armananzas, R. Win-win data sharing in neuroscience. *Nat. Methods* **14**, 112–116 (2017).
54. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans. Phil. Trans. R. Soc. Lond. B* **314**, 1–340 (1986).
55. Denk, W. & Horstmann, H. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol.* **2**, e329 (2004).
A seminal paper of the modern age of ultrastructural connectomics, this study introduced the high-throughput serial scanning electron microscopy method that remains the most used in research (see REFS 56 and 57 for other methods).
56. Knott, G., Marchman, H., Wall, D. & Lich, B. Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *J. Neurosci.* **28**, 2959–2964 (2008).
57. Hayworth, K. J. *et al.* Imaging ATUM ultrathin section libraries with WaferMapper: a multi-scale approach to EM reconstruction of neural circuits. *Front. Neural Circuits* **8**, 68 (2014).
58. Kasthuri, N. *et al.* Saturated reconstruction of a volume of neocortex. *Cell* **162**, 648–661 (2015).
59. Bock, D. D. *et al.* Network anatomy and *in vivo* physiology of visual cortical neurons. *Nature* **471**, 177–182 (2011).
60. Eberle, A. L. *et al.* High-resolution, high-throughput imaging with a multibeam scanning electron microscope. *J. Microsc.* **259**, 114–120 (2015).
61. Beier, T. *et al.* Multicut brings automated neurite segmentation closer to human performance. *Nat. Methods* **14**, 101–102 (2017).
62. Greene, M. J., Kim, J. S., Seung, H. S. & The EyeWirers. Analogous convergence of sustained and transient inputs parallel on off pathways retinal motion computation. *Cell Rep.* **14**, 1892–1900 (2016).
63. Kim, J. S. *et al.* Space-time wiring specificity supports direction selectivity in the retina. *Nature* **509**, 331–336 (2014).
References 62 and 63 used serial block-face imaging (described in Ref 55) and crowd-sourced reconstruction to provide an ultrastructural classification of retinal BCs.
64. Berck, M. E. *et al.* The wiring diagram of a glomerular olfactory system. *eLife* **5**, e14859 (2016).
65. Takemura, S. Y. *et al.* A visual motion detection circuit suggested by *Drosophila* connectomics. *Nature* **500**, 175–181 (2013).
66. Briggman, K. L., Helmstaedter, M. & Denk, W. Wiring specificity in the direction-selectivity circuit of the retina. *Nature* **471**, 183–188 (2011).
67. Helmstaedter, M. *et al.* Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature* **500**, 168–174 (2013).
68. Lee, W. C. *et al.* Anatomy and function of an excitatory network in the visual cortex. *Nature* **532**, 370–374 (2016).
69. Morgan, J. L., Berger, D. R., Wetzel, A. W. & Lichtman, J. W. The fuzzy logic of network connectivity in mouse visual thalamus. *Cell* **165**, 192–206 (2016).
70. Obien, M. E., Deligkaris, K., Bullmann, T., Bakkum, D. J. & Frey, U. Revealing neuronal function through microelectrode array recordings. *Front. Neurosci.* **8**, 423 (2014).
71. Antic, S. D., Empson, R. M. & Knopfel, T. Voltage imaging to understand connections and functions of neuronal circuits. *J. Neurophysiol.* **116**, 135–152 (2016).
72. Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
73. Peron, S., Chen, T. W. & Svoboda, K. Comprehensive imaging of cortical networks. *Curr. Opin. Neurobiol.* **32**, 115–123 (2015).
74. Yang, W. & Yuste, R. *In vivo* imaging of neural activity. *Nat. Methods* **14**, 349–359 (2017).
75. Kim, T. H. *et al.* Long-term optical access to an estimated one million neurons in the live mouse cortex. *Cell Rep.* **17**, 3385–3394 (2016).
76. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for *in vivo* imaging. *eLife* **5**, e14472 (2016).
77. Andermann, M. L. *et al.* Chronic cellular imaging of entire cortical columns in awake mice using microprisms. *Neuron* **80**, 900–913 (2013).
78. Ghosh, K. K. *et al.* Miniaturized integration of a fluorescence microscope. *Nat. Methods* **8**, 871–878 (2011).
79. Lu, R. *et al.* Video-rate volumetric functional imaging of the brain at synaptic resolution. *Nat. Neurosci.* **20**, 620–628 (2017).
80. Ouzounov, D. G. *et al. In vivo* three-photon imaging of activity of GCaMP6-labeled neurons deep in intact mouse brain. *Nat. Methods* **14**, 388–390 (2017).
81. Ziv, Y. *et al.* Long-term dynamics of CA1 hippocampal place codes. *Nat. Neurosci.* **16**, 264–266 (2013).
82. Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M. & Keller, P. J. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods* **10**, 413–420 (2013).
83. Naumann, E. A. *et al.* From whole-brain data to functional circuit models: the zebrafish optomotor response. *Cell* **167**, 947–960e20 (2016).
84. Cox, J., Pinto, L. & Dan, Y. Calcium imaging of sleep-wake related neuronal activity in the dorsal pons. *Nat. Commun.* **7**, 10763 (2016).
85. Resendez, S. L. *et al.* Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. *Nat. Protoc.* **11**, 566–597 (2016).
86. Shapiro, E., Biezuner, T. & Linnarsson, S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat. Rev. Genet.* **14**, 618–630 (2013).
87. Trapnell, C. Defining cell types and states with single-cell genomics. *Genome Res.* **25**, 1491–1498 (2015).
88. Poulin, J. F., Tasic, B., Hjerling-Leffler, J., Trimarchi, J. M. & Awatramani, R. Disentangling neural cell diversity using single-cell transcriptomics. *Nat. Neurosci.* **19**, 1131–1141 (2016).
89. Svensson, V. *et al.* Power analysis of single-cell RNA-sequencing experiments. *Nat. Methods* **14**, 381–387 (2017).
90. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).
91. Ramskold, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30**, 777–782 (2012).
92. Islam, S. *et al.* Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* **21**, 1160–1167 (2011).

93. Islam, S. *et al.* Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Methods* **11**, 163–166 (2014).
94. Hashimshony, T. *et al.* CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, 77 (2016).
95. Hashimshony, T., Wagner, F., Sher, N. & Yanai, I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* **2**, 666–673 (2012).
96. Pollen, A. A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).
97. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
This paper introduces Drop-seq, one of three similar methods used for microfluidic-based high-throughput scRNA-seq (see also REFS 98 and 99).
98. Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
99. Zheng, G. X. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
100. Sheng, K., Cao, W., Niu, Y., Deng, Q. & Zong, C. Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat. Methods* **14**, 267–270 (2017).
101. Habib, N. *et al.* Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* **353**, 925–928 (2016).
102. Krishnaswami, S. R. *et al.* Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat. Protoc.* **11**, 499–524 (2016).
103. Lacar, B. *et al.* Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat. Commun.* **7**, 11022 (2016).
104. Lake, B. B. *et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
105. Shekhar, K. *et al.* Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323e30 (2016).
This paper uses high-throughput scRNA-seq (Drop-seq, REF. 97) to provide a molecular classification of retinal BCs; it also shows a 1:1 relationship of molecularly and morphologically defined types.
106. Ziegenhain, C. *et al.* Comparative analysis of single-cell RNA sequencing methods. *Mol. Cell* **65**, 631–643e4 (2017).
107. Cadwell, C. R. *et al.* Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nat. Biotechnol.* **34**, 199–203 (2016).
108. Fuzik, J. *et al.* Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat. Biotechnol.* **34**, 175–183 (2016).
109. Baden, T. *et al.* The functional diversity of retinal ganglion cells in the mouse. *Nature* **529**, 345–350 (2016).
110. Lubeck, E. & Cai, L. Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat. Methods* **9**, 743–748 (2012).
111. Shah, S., Lubeck, E., Zhou, W. & Cai, L. *In situ* transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* **92**, 342–357 (2016).
112. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
This paper describes MERFISH, one of several highly multiplexed methods used for *in situ* hybridization.
113. Moffitt, J. R. *et al.* High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence *in situ* hybridization. *Proc. Natl Acad. Sci. USA* **113**, 11046–11051 (2016).
114. Moffitt, J. R. *et al.* High-performance multiplexed fluorescence *in situ* hybridization in culture and tissue with matrix imprinting and clearing. *Proc. Natl Acad. Sci. USA* **113**, 14456–14461 (2016).
115. Chen, F. *et al.* Nanoscale imaging of RNA with expansion microscopy. *Nat. Methods* **13**, 679–684 (2016).
116. Lee, J. H. *et al.* Highly multiplexed subcellular RNA sequencing *in situ*. *Science* **343**, 1360–1363 (2014).
117. Ke, R. *et al.* *In situ* sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10**, 857–860 (2013).
118. Larsson, C., Grundberg, I., Soderberg, O. & Nilsson, M. *In situ* detection and genotyping of individual mRNA molecules. *Nat. Methods* **7**, 395–397 (2010).
119. Masland, R. H. The neuronal organization of the retina. *Neuron* **76**, 266–280 (2012).
120. Sanes, J. R. & Zipursky, S. L. Design principles of insect and vertebrate visual systems. *Neuron* **66**, 15–36 (2010).
121. Kay, J. N., Chu, M. W. & Sanes, J. R. MEGF10 and MEGF11 mediate homotypic interactions required for mosaic spacing of retinal neurons. *Nature* **483**, 465–469 (2012).
122. Reese, B. E. & Keeley, P. W. Design principles and developmental mechanisms underlying retinal mosaics. *Biol. Rev. Camb. Philos. Soc.* **90**, 854–876 (2015).
123. Euler, T., Haverkamp, S., Schubert, T. & Baden, T. Retinal bipolar cells: elementary building blocks of vision. *Nat. Rev. Neurosci.* **15**, 507–519 (2014).
124. Franke, K. *et al.* Inhibition decorrelates visual feature representations in the inner retina. *Nature* **542**, 439–444 (2017).
This paper uses high-throughput calcium imaging to provide a physiological classification of retinal BCs; the types defined here correspond 1:1 with those defined molecularly (REF. 105) and morphologically (REFS 62 and 63).
125. Wassle, H., Puller, C., Muller, F. & Haverkamp, S. Cone contacts, mosaics, and territories of bipolar cells in the mouse retina. *J. Neurosci.* **29**, 106–117 (2009).
126. Della Santina, L. *et al.* Glutamatergic monopolar interneurons provide a novel pathway of excitation in the mouse retina. *Curr. Biol.* **26**, 2070–2077 (2016).
127. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
128. Douglas, R. J. & Martin, K. A. Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* **27**, 419–451 (2004).
129. Greig, L. C., Woodworth, M. B., Galazo, M. J., Padmanabhan, H. & Macklis, J. D. Molecular logic of neocortical projection neuron specification, development and diversity. *Nat. Rev. Neurosci.* **14**, 755–769 (2013).
130. Harris, K. D. & Shepherd, G. M. The neocortical circuit: themes and variations. *Nat. Neurosci.* **18**, 170–181 (2015).
An updated review of neocortical neuronal types and their patterns of input-output connections, which are repeated across cortical areas.
131. Molyneaux, B. J., Arlotta, P., Menezes, J. R. & Macklis, J. D. Neuronal subtype specification in the cerebral cortex. *Nat. Rev. Neurosci.* **8**, 427–437 (2007).
132. Kim, E. J., Juavinett, A. L., Kyubwa, E. M., Jacobs, M. W. & Callaway, E. M. Three types of cortical layer 5 neurons that differ in brain-wide connectivity and function. *Neuron* **88**, 1253–1267 (2015).
133. Kita, T. & Kita, H. The subthalamic nucleus is one of multiple innervation sites for long-range corticofugal axons: a single-axon tracing study in the rat. *J. Neurosci.* **32**, 5990–5999 (2012).
134. Velez-Fort, M. *et al.* The stimulus selectivity and connectivity of layer six principal cells reveals cortical microcircuits underlying visual processing. *Neuron* **83**, 1431–1443 (2014).
135. Yamashita, T. *et al.* Membrane potential dynamics of neocortical projection neurons driving target-specific signals. *Neuron* **80**, 1477–1490 (2013).
136. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
This article reports a comprehensive scRNA-seq characterization and classification of adult neocortical neurons using the Smart-seq method, which resulted in a transcriptomic cell type taxonomy that is supported by genetic (Cre recombinase driver lines), physiological, morphological and projectional evidence.
137. Fishell, G. & Rudy, B. Mechanisms of inhibition within the telencephalon: "where the wild things are". *Annu. Rev. Neurosci.* **34**, 535–567 (2011).
138. Kepecs, A. & Fishell, G. Interneuron cell types are fit to function. *Nature* **505**, 318–326 (2014).
139. Rudy, B., Fishell, G., Lee, S. & Hjerling-Leffler, J. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev. Neurobiol.* **71**, 45–61 (2011).
140. Tremblay, R., Lee, S. & Rudy, B. GABAergic interneurons in the neocortex: from cellular properties to circuits. *Neuron* **91**, 260–292 (2016).
An updated review of neocortical interneuron types, their cellular properties and their potential functions in circuit motifs and network operations.
141. Klausberger, T. & Somogyi, P. Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. *Science* **321**, 53–57 (2008).
142. Jiang, X. *et al.* Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, aac9462 (2015).
Reports a large-scale electrophysiological and morphological characterization of adult neocortical neurons, revealing connectivity patterns among morphologically defined neuronal types.
143. Ohki, K. & Reid, R. C. Specificity and randomness in the visual cortex. *Curr. Opin. Neurobiol.* **17**, 401–407 (2007).
144. Morrie, R. D. & Feller, M. B. Development of synaptic connectivity in the retinal direction selective circuit. *Curr. Opin. Neurobiol.* **40**, 45–52 (2016).
145. Oyster, C. W. & Barlow, H. B. Direction-selective units in rabbit retina: distribution of preferred directions. *Science* **155**, 841–842 (1967).
146. Reid, R. C. From functional architecture to functional connectomics. *Neuron* **75**, 209–217 (2012).
147. Dehorter, N. *et al.* Tuning of fast-spiking interneuron properties by an activity-dependent transcriptional switch. *Science* **349**, 1216–1220 (2015).
148. Spitzer, N. C. Neurotransmitter switching? No surprise. *Neuron* **86**, 1131–1144 (2015).
149. Gray, K. A., Seal, R. L., Tweedie, S., Wright, M. W. & Bruford, E. A. A review of the new HGNC gene family resource. *Hum. Genom.* **10**, 6 (2016).
150. McDonald, A. G., Boyce, S. & Tipton, K. F. ExplorEnz: the primary source of the IUBMB enzyme list. *Nucleic Acids Res.* **37**, D593–D597 (2009).
151. Hamann, J. *et al.* International Union of Basic and Clinical Pharmacology. XCIV. Adhesion G protein-coupled receptors. *Pharmacol. Rev.* **67**, 338–367 (2015).
152. Usoskin, D. *et al.* Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat. Neurosci.* **18**, 145–153 (2015).
153. Zeisel, A. *et al.* Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
154. Darmanis, S. *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl Acad. Sci. USA* **112**, 7285–7290 (2015).
155. Pollen, A. A. *et al.* Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015).
156. Romanov, R. A. *et al.* Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nat. Neurosci.* **20**, 176–188 (2017).
157. Li, C. L. *et al.* Somatosensory neuron types identified by high-coverage single-cell RNA-sequencing and functional heterogeneity. *Cell Res.* **26**, 83–102 (2016).
158. Kee, N. *et al.* Single-cell analysis reveals a close relationship between differentiating dopamine and subthalamic nucleus neuronal lineages. *Cell Stem Cell* **20**, 29–40 (2017).
159. Campbell, J. N. *et al.* A molecular census of arcuate hypothalamus and median eminence cell types. *Nat. Neurosci.* **20**, 484–496 (2017).
160. Wallace, M. L. *et al.* Genetically distinct parallel pathways in the entopeduncular nucleus for limbic and sensorimotor output of the basal ganglia. *Neuron* **94**, 138–152e5 (2017).
161. Quadrato, G. *et al.* Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* **545**, 48–53 (2017).

Acknowledgements
We thank our colleagues K. Shekhar, L. Zipursky, E. Lein, C. Reid, M. Hawrylycz, B. Tasic, S. Sorensen, J. Berg and C. Koch for their valuable comments. This work was funded by National Institutes of Health grants R37NS029169 and U01MH105960 to J.R.S. and U01MH105982 to H.Z. H.Z. would like to thank the Allen Institute founder, Paul G. Allen, for his vision, encouragement and support.

Competing interests statement
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regards to jurisdictional claims in published maps and institutional affiliations.

FURTHER INFORMATION
NeuroMorpho.org: http://neuromorpho.org/
Eyewire: http://museum.eyewire.org
Allen Cell Types Database: http://celltypes.brain-map.org/

ALL LINKS ARE ACTIVE IN THE ONLINE PDF